Applying AD to the Community Land Model

Laboratory for Advanced Numerical Simulations
Mathematics and Computer Science
Argonne National Laboratory

Azamat Mametjanov, **Boyana Norris**,
Xiaoyan Zeng, Beth Drewniak,
Jean Utke, Mihai Anitescu, and **Paul Hovland**

Supported by DOE's BER under Climate Science for Sustainable Energy Future (CSSEF)

Outline

CESM and CLM

CESM: Community Earth System Model

- Is a global model for simulations of the climate system.
- Is composed of five fully-coupled components of
  1. atmosphere
  2. *land*
  3. ocean
  4. land-ice
  5. sea-ice
- Each component is configurable for one the modes
  - active (fully prognostic)
  - data (intercomponent data cycling)
  - inactive (interface)
- Provides state-of-the-art simulations of past, present and future climate states (1850–2100)

CLM: Community Land Model

Is a CESM component for simulations of

1. energy fluxes in land bio-geophysics
2. chemical compound fluxes in land bio-geochemistry
3. water fluxes in land hydrology

Divides the modeled surface grid into grid-cells

- land units
  - columns
    - plant functional types (PFT)

CLM: Structure

Land unit models

- glacier
- lake
- wetland
- urban area
- vegetated area

Column models vertical layers of

- soil – up to 15 layers
- snow – up to 5 layers

Plant functional type models

- trees and shrubs – 11 types
- grasses – 3 types
- crops – 1 type

CLM extension: CLM-Crop

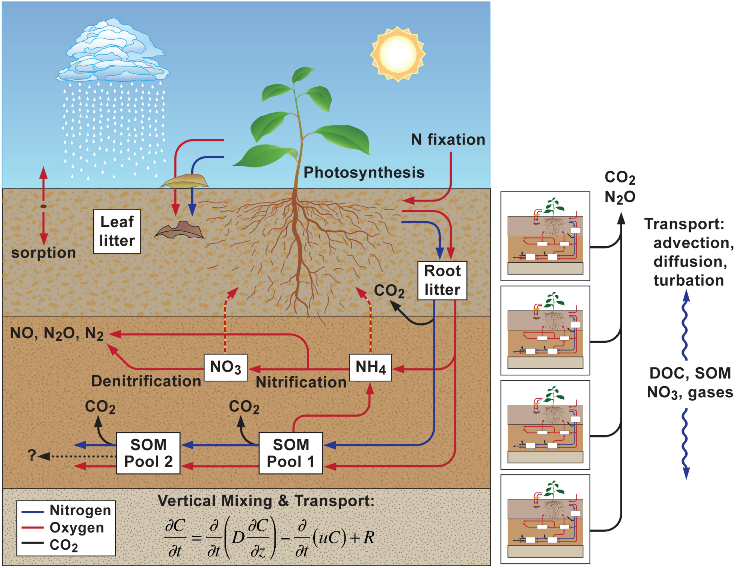
Extends CLM with irrigated/managed crop species

- corn
- wheat
- soybean

Models pools and fluxes of Carbon and Nitrogen (CN) in

- leaves
- stems
- roots
- harvested organs

CLM-Crop



N fixation
Photosynthesis
Leaf litter
sorption
Root litter
CO_2
NO, N_2O, N_2
NO_3
NH_4
Denitrification
Nitrification
CO_2
CO_2
SOM Pool 2
SOM Pool 1
?
Nitrogen
Oxygen
CO_2
Vertical Mixing & Transport:
$\frac{\partial C}{\partial t} = \frac{\partial}{\partial t}\left(D\frac{\partial C}{\partial z}\right) - \frac{\partial}{\partial t}(uC) + R$
CO_2
N_2O
Transport: advection, diffusion, turbation
DOC, SOM
NO_3, gases

CLM-Crop Parameter Calibration

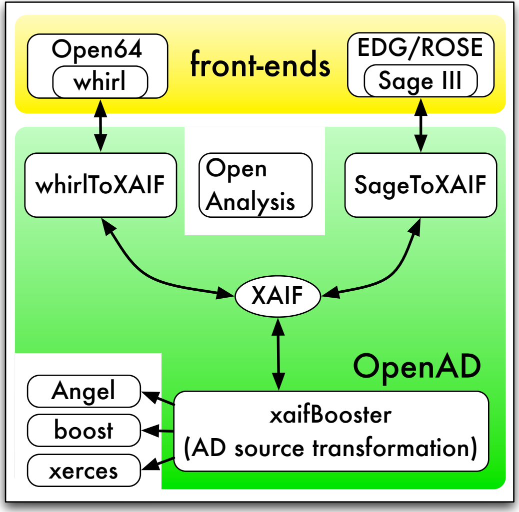
- Crop growth is substantially affected by CN fluxes
- CN fluxes are determined by CN ratios of a given crop species
- There is a large uncertainty in regards to CN ratios
- We study sensitivity of CN ratios to calibrate the parameters for further improvement of model accuracy

AD and OpenAD

Algorithmic Differentiation

- Computation encoded by a program can be broken into a sequence of elemental operations
- Independent (and *varied*) variable values are transformed into (*useful* and) dependent variable values
- AD computes derivatives of elemental operations acting on active variables
- Based on the chain rule, derivative values are accumulated to obtain a derivative of the entire program

OpenAD: modular Fortran/C/C++ AD tool

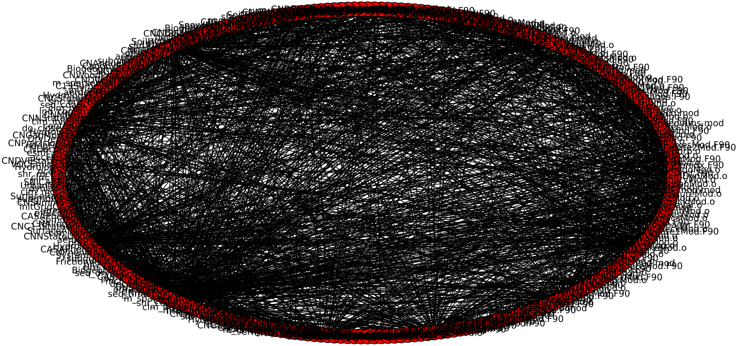


AD Process

CLM source code

Modules	Files	Lines
Biogeochemistry	39	18568
Biogeophysics	24	13931
Riverroute	3	1435
Couplers/drivers	53	34416
Total	119	68350

CLM module dependencies



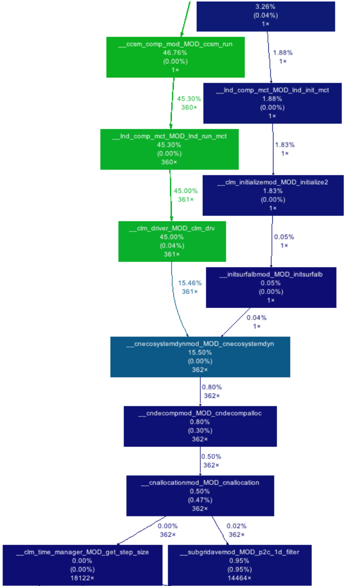
Call graphs

- Help to understand control flow dependencies
- CESM is supported by PGI complier suite
- Ported to GNU suite to profile with `gprof`

AD pattern: start small, iterative increments

- AD is excellent when applied to numeric programming of smooth functions
- Real codes represent complex compositions of (non-smooth) computations
- Pattern
  - start with a small numerical core C
  - differentiate: C'
  - validate C'
  - add more functions/subroutines
  - repeat until the desired scope is achieved

Local AD: CNAllocation module



OpenAD work-flow

- Identify module dependencies
- Collect source files
- Preprocess with proper build sequence parameters
- Concatenate the sources into a single file
- Insert annotations marking independent and dependent variables
- Invoke transformations:
  1. preprocess
  2. fortran$\rightarrow$whirl
  3. whirl$\rightarrow$xaif
  4. xaif$\rightarrow$xaif'
  5. xaif'$\rightarrow$whirl'
  6. whirl'$\rightarrow$fortran'
  7. postprocess

Results

AD of CNAllocation calculation

12 inputs

- final leaf/stem/root CN ratio
- leaf/organ/live-wood/dead-wood CN ratio
- new fine-root/stem C per new leaf C
- new coarse-root C per new stem C
- fraction of new wood that is live
- fraction of allocation that goes to current growth

6 outputs

- leaf/stem/organ carbon/nitrogen

Select derivative results

	CORN	
	C	*N*
LEAF		
fleafcn	7.0353917	-93.0305059
frootcn	4.6136744	-46.2578484
fstemcn	1.9315305	0.1931531
deadwdcn	0.0000000	0.0000000
STEM		
fleafcn	14.4554170	0.2891083
frootcn	9.4894822	-12.9865931
fstemcn	3.9686602	-25.9019159
deadwdcn	0.0000000	0.0000000
ORGAN		
fleafcn	1277.0876953	25.5417539
frootcn	809.0263835	16.1805277
fstemcn	350.6178413	7.0123568
deadwdcn	0.0000000	0.0000000

Lessons Learned

CESM application

- Highly structured, hierarchical Fortran 90 code
- State preservation via global variables
- Highly configurable (`ifdef`) code

Need greater tool support

- Control flow graph: dependency analysis
- Data flow graph: activity analysis
- Automated active variable type-changing

Thank you